

## Method to process two audio input signals

The invention relates to a method for the processing of at least two input signals which contain audio information and possibly also video information, in which method the audio information and possibly also video information of a first input signal is processed for acoustic and possibly also audiovisual reproduction.

5           The invention also relates to a device for the processing of at least two input signals which contain audio information and possibly also video information, which device comprises a reproduction device for the reproduction of a first input signal.

10           It is known to provide television signals with text, in addition to the audio and video information of a television program, which text contains, for example, headlines, stock exchange data or other current information. It is also known to reproduce a second television signal optically in a small section of the display screen. The audio signal of this further television signal in the so-called PIP (picture-in-picture) method is not reproduced. Also  
15           known are inserted texts which optically reproduce the audio signal of the reproduced television signal at least partly for the benefit of persons who are deaf or hard of hearing.

          United States patent 5,557,338 A discloses a television system in which the picture comprises a main picture and a secondary picture and in which additionally text information in the form of a subtitle is reproduced in the main picture, which text information  
20           relates to the broadcast reproduced in the secondary image. The transmitter then has to transmit the text information together with the information of the secondary picture. This system constitutes an extension of the so-called PIP (picture-in-picture) method in which text information is reproduced in addition to the secondary picture.

25

          It is an object of the present invention to provide a method and a device of the kind set forth whereby at least one further input signal can be reproduced in addition to a reproduced input signal. The reception of at least one further acoustic or audiovisual input signal is thus made possible wherever an acoustic or audiovisual input signal is already

received. It should be possible to use the method also in locations where acoustic reception of an input signal is not possible, for example, because of excessive ambient noise.

In respect of the method the object in accordance with the invention is achieved by means of a method for the processing of at least two input signals which contain  
5 audio information and possibly also video information, in which method the audio information and possibly also the video information of the one input signal is processed for acoustic and possibly also audiovisual reproduction, at least one second input signal is applied to speech recognition means, text information concerning the audio information contained in at least the second input signal is determined by means of the speech recognition  
10 means, and the text information determined is optically reproduced.

The method in accordance with the invention thus enables different input signals to be processed in such a manner that the speech occurring therein is recognized and converted into text which is optically reproduced. This enables, for example, the text of a different television broadcast to be inserted in the picture during the reception of a television  
15 broadcast. The user can thus be informed about other topics during the reception of a television broadcast. The input signal whose speech is recognized may then also originate from a different external source, for example, from a radio receiver, a video recorder or also from a telephone line. The information received in the form of an audio signal from a radio station can thus be reproduced as text during the reproduction of a television broadcast. It is  
20 also possible to optically reproduce incoming telephone calls which are routed to a telephone answering machine, so that the user can obtain information concerning the call and, for example, decide whether or not to accept the call. The speech recognition makes it possible to process practically any input signal containing audio information and possibly also video information and to reproduce such an input signal in addition to a first input signal.

25 The object in accordance with the invention is also achieved by means of a device for the processing of at least two input signals which contain audio information and possibly also video information, which device comprises a reproduction device for the reproduction of an input signal, speech recognition means for determining text information contained in the audio information of at least one second input signal, and an optical  
30 reproduction device for the reproduction of the text information determined.

The speech recognition means may be separate from the reproduction device of the one input signal and the optical reproduction device for the reproduction of the text information determined, or be integrated in one of said devices. It is also possible for all components of the device in accordance with the invention to be integrated in one apparatus,

for example, in a television receiver. The external or integrated speech recognition means enable the processing of the audio information of at least one second input signal and to optically reproduce the text information determined therefrom in addition to a first input signal.

5           The text information is advantageously reproduced as a running text, the speed of the running text being automatically adapted to the reproduction. It is also possible to buffer the text information and to reproduce it in a delayed fashion. For example, a radio broadcast could be processed at predetermined instants by means of speech recognition means, and the text information determined, for example, the headlines, could be buffered  
10           and be optically reproduced at predetermined instants, or at instants selected by the user, during the reproduction of an input signal.

          The video information of the one input signal and the text information of the at least one further input signal are advantageously reproduced on a common monitor. If the first input signal reproduced is not a video signal, the text information of the at least one  
15           further input signal can be reproduced on a suitable display which is provided especially for this purpose or which is already present. For example, the first input signal may be the acoustic signal of a telephone and a second incoming telephone call can be optically reproduced on the display of the telephone.

          The second input signal can advantageously be selected by the user. The user  
20           can thus decide which text information is additionally reproduced in an optical fashion during the reproduction of an input signal.

          The selection of the second input signal can then be performed on the basis of stored information. This information may involve given criteria as selected by the user or may also concern automatically detected user habits.

25           Parameters of the speech recognition means are advantageously modified on the basis of the text information of the second input signal. As a result, for example, the speech recognition means can be optimally adapted to the second input signal in that, for example, appropriate libraries or languages adapted to the second input signal are selected by recognition of given texts.

30           It is also advantageous when the text information determined is compared with stored texts and given steps are taken when given comparison results are obtained. For example, the optical reproduction of the text information can be rendered dependent on the correspondence with stored texts. As a result of this feature, it is possible to insert the text

only subject to given conditions. In this respect, for example, given keywords can be used as a criterion.

Additionally it may be arranged that in the case of correspondence between the text information and given stored texts the audio information and possibly also video information of the second input signal is reproduced instead of the audio information and possibly also video information of the first input signal. For example, the at least one further input signal can thus be monitored so that automatic switching over to this input signal can take place, for example, at the beginning of a news broadcast or at the beginning of a sports broadcast.

The input signals to be reproduced are advantageously television signals. However, various other input signals, for example, radio signals, telephone signals or the like, are also feasible.

The reproduction device for the reproduction of an input signal and the reproduction device for the reproduction of the text information determined are advantageously formed by a common monitor.

When storage means are provided for the storage of the text information determined, the text information contained in the audio information of at least one further input signal can be stored for later or repeated reproduction.

In order to enable the user to choose from among a plurality of input signals available, in conformity with a further feature of the invention there are provided control means. Such control means may be connected to a memory for information, so that the selection of the at least one second input signal can take place on the basis of the information stored in the memory.

When a switching device is provided for switching over parameters of the speech recognition means, optimum adaptation of the speech recognition means can be achieved on the basis of the text information of the second input signal. For example, upon recognition of the language of the second input signal, the speech recognition means can be adapted to this language and the relevant libraries can be activated.

Advantageously there is provided a comparison unit for comparing the text information with stored texts. This offers a series of further options, for example, text-dependent reproduction of the text information or the like.

In order to enable text-specific reproduction of the text information of a second input signal, said comparison unit may be connected to the optical reproduction unit.

Furthermore, there may be provided a switching unit for switching over the reproduction of the input signals; such a switching unit is connected to the comparison unit. The switching unit may then be formed by said control means for the selection of the input signals.

5                   The reproduction device for the reproduction of an input signal may be formed by a television receiver.

Embodiments of the invention will be described in detail hereinafter with  
10 reference to the drawings, however, without the invention being restricted thereto in any way.

Fig. 1 shows a block diagram of an embodiment of the device for the processing of at least two input signals which contain audio information and possibly also video information.

Fig. 2 shows an example of the reproduction devices for the input signal and  
15 the text information determined.

Fig. 3 shows an extended block diagram of a device in accordance with the invention.

Fig. 4 shows an example of an application in the form of a master control room.

20                   Fig. 5 shows a further application concerning a telephone set.

Fig. 1 shows a block diagram of a device for the processing of at least two input signals  $S_i$  which contain audio information  $A_i$  and possibly also video information  $V_i$ .  
25 The device shown serves to process two input signals  $S_1, S_2$ , but can be extended at will to an arbitrary number of input signals  $S_i$ . The device includes a reproduction device 10 for the reproduction of an input signal  $S_1$ , for example, a television receiver, which processes and reproduces the audio information  $A_1$  and possibly also video information  $V_1$  of the input signal  $S_1$ . The at least one second input signal  $S_2$  is applied to speech recognition means 11 in  
30 which the text information  $T_2$  which is contained in the audio information  $A_2$  of the input signal  $S_2$  is determined. This text information  $T_2$  is reproduced by means of an optical reproduction device 12. It is thus possible to reproduce, in addition to the input signal  $S_1$ , also the text information  $T_2$  contained in a further input signal  $S_2$ , that is, simultaneously or shifted in time. In order to enable time-shifted reproduction there may be provided storage

means 14 for the storage of the text information  $T_2$  determined. Depending on the type of input signal  $S_1$ ,  $S_2$ , it may be advantageous to integrate the reproduction device 10 for the reproduction of the input signal  $S_1$  and the reproduction device 12 for the reproduction of the text information  $T_2$  determined in a common monitor 13 or the like.

5 Fig. 2 shows an example of such a common monitor 13 which comprises the reproduction device 10 for the reproduction of the first input signal  $S_1$ , for example, a television broadcast, and also the optical reproduction device 12 for the text information  $T_2$  determined. The text information  $T_2$  is thus inserted in the form of subtitles in the television picture of the input signal  $S_1$ .

10 Fig. 3 shows a block diagram of a device for the processing of a plurality of input signals  $S_i$  which has been extended in comparison with that shown in Fig. 1. A plurality of input signals  $S_i$  which contain audio information  $A_i$  and possibly also video information  $V_i$  is applied to control means 15 which serve for the selection of the input signals  $S_i$ . A first input signal  $S_1$  is then suitably processed and reproduced on a reproduction device 10. At  
15 least one further input signal  $S_2$  is applied to the speech recognition means 11 and the text information  $T_2$  which is contained in the audio information  $A_2$  of the input signal  $S_2$  is determined therefrom. The text information  $T_2$  may be applied to a switching device 17 for switching over parameters  $P_i$  of the speech recognition means 11, thus enabling optimum adaptation of the speech recognition means 11 to the processed text information  $T_2$ . In  
20 addition, the text information  $T_2$  can be applied to a comparison unit 18 prior to the optical reproduction, the text information  $T_2$  then being compared with texts  $T_s$  which are stored in a memory 19 in said comparison unit. As a result of this comparison in the comparison unit 18, for example, text-specific reproduction of the text information  $T_2$  can take place on the optical reproduction device 12. Moreover, the comparison unit 18 may be connected to the  
25 control means 15 or to a further switching unit (not shown) so that when a given stored text  $T_s$  is recognized in the text information  $T_2$ , switching over to a different input signal  $S_i$  may take place. A memory 16 can serve for the storage of information  $I_i$  which may concern, for example, given user habits. The memory 16 is advantageously connected to the control means 15 so that selection of the input signals  $S_i$  can be carried out on the basis of the  
30 information  $I_i$  stored in the memory 16. The reproduction device 10 for the reproduction of an input signal  $S_1$  and the optical reproduction device 12 for the reproduction of the text information  $T_2$  determined can be integrated in a common monitor 13. Moreover, all of the devices in accordance with the invention may be integrated in one apparatus, for example, a television receiver 20.

Fig. 4 shows an application of the invention for a master control room in which, by way of example, a plurality of monitors 21 is provided for the reproduction of video information  $V_1$  to  $V_8$  and audio signals  $A_1$  to  $A_8$  of eight input signals  $S_1$  to  $S_8$ . Each time only one audio signal  $A_i$  can be received. The other audio signals  $A_i$  of the input signals  $S_i$  or audio signals from other sources, for example, the audio signals from the camera men or the associated sound technicians, can be displayed on the monitors 21 in the form of text information  $T_1$  to  $T_8$ , thus providing the director with further information for the selection of the signal  $S_i$  to be broadcast.

Fig. 5 shows a further application of the invention in a telephone set 22, in which, during the reception of a telephone call, the text information  $T_2$  of a further telephone call can be displayed additionally on an optical display device 12 in the form of a display customarily provided in telephone sets. The invention thus offers the user of the telephone set 22 the simultaneous reception of a further telephone call which is diverted, for example, to a telephone answering apparatus. For example, the user can then decide to interrupt the first telephone call and switch over to the second telephone call.

The present invention is by no means restricted to the described examples and can also be applied to various other input signals.